Math. Program., Ser. B (2001)

Digital Object Identifier (DOI) 10.1007/s101070100258

B.T. Polyak

History of mathematical programming in the USSR: analyzing the phenomenon*

Received: January 29, 2001 / Accepted: May 17, 2001

Abstract. I am not a historian; these are just reminiscences of a person involved in the development of optimization theory and methods in the former USSR. I realize that my point of view may be very personal; however, I am trying to present as broad and unbiased picture as I can.

1. Mathematics in the USSR

Let me start with a few words about the general situation in mathematics and natural science in the USSR. For a younger reader who lives nowadays in a completely different environment, these remarks are worth reading, since today it is really hard to imagine the rather specific life conditions in that country in the 1950–60s. There was a communist totalitarian system with a very poor population and an obsolete industry. No human rights were respected, no basic liberties like freedom of speech were available. Nevertheless, certain branches of science (such as mathematics) were flourishing, and, in some areas, the achievements in research were comparable or even exceeded those of the rest of the world. What were the reasons for such a success?

First, there was strong government support for research in such fields as physics and mathematics. This can be explained mainly as motivated by military reasons. The authorities were convinced that success in atomic bomb construction, the space competition, etc., can be achieved via the development of the underlying scientific disciplines. Hence, fundamental research was supported as well as applied studies. Mathematicians and physicists basically had no problems with funding. There were many research institutions not related to education, where the employees could concentrate on purely theoretical studies, with minor limitations on the duration of the research and having no strict obligations on the results obtained. Such freedom of research, in combination with tenure positions for scientists of all levels, sometimes led to outstanding results.

Second, there were good traditions in teaching mathematics. Often, the level of mathematical education in Soviet high schools was superior to that of the USA. Moreover, for talented middle- and high-school students, there was a specific system of mathematical training based on mathematical circles supervised by University students and professors. Every year, mathematical olympiads of various levels were organized.

B.T. Polyak: Institute for Control Science, Moscow, Russia

* Plenary talk at 17th International Symposium on Mathematical Programming, Atlanta, GA, Aug. 7–11, 2000

I always recall one such circle (within Moscow State University) and the Moscow mathematical olympiads with delight; they shaped my professional future. Mathematical departments at some Universities were highly remarkable. For instance, at the mathematics department of Moscow State University in the 1960s, the concentration of outstanding mathematicians of the XX-th century had no analogs in the rest of the world (A. Kolmogorov, I. Gelfand, L. Pontryagin, P. Aleksandrov, A. Tikhonov, L. Ljusternik, V. Arnold, Yu. Manin, S. Novikov, Ya. Sinai, R. Dobrushin and many others). There was a system of seminars chaired by such scientists, and many students participated in these seminars from their undergraduate years. It was an exceptional opportunity to contact such legendary figures and to be involved in research at the early stages.

Third, tuition at universities was free, and in principle, a student of modest means from a provincial town was admitted to learn at the best, Moscow University, provided he/she passed the entrance examinations. Textbooks and monographs were extremely cheap and were published in a large number of copies. For instance, the price of the textbook "Mathematical Programming" by V. Karmanov was 44 kopecks (about 70 cents), and just for one edition the circulation was 60,000 copies; the price of a popular-science book "Stories on Maxima and Minima" by V. Tikhomirov was 35 kopecks with a circulation of 160,000 copies [1,2].

One more reason: there was a deficiency of career opportunities for ambitious young people. Recall that there were no such professions as a businessman, a manager, a banker, or a programmer. It was impossible to pursue the career of a politician, a judge or a diplomat without being a member of the Communist party. Moreover, there was no hope to perform any honest research in most humanitarian sciences, which were strongly politicized and filled with Marxist terminology. Thus the career of a mathematician was one of the few ones free of ideological pressure.

However, this idyllic picture had a dark side. "The iron curtain" was not only a metaphor, it was a real obstacle to international contacts. Academician N.N. Lusin was exposed to a humiliating execution after the publication of his papers in Western journals in 1936 [3]. When Professor Ya.Z. Tsypkin received a letter in the late 1940s from an American reader of his paper, he was summoned by the KGB and underwent a long investigation there, tottering at the edge of arrest. Of course, the situation in the 1950–60s was not that dramatic, but lots of difficulties still remained. Even if a scientist was invited to a conference abroad, with full coverage of his expenses, this did not mean anything. First, there was the "black list" of scientists who were not allowed to go abroad under any circumstances. The reason could be his dissident activity (e.g., he had signed a protest against putting the mathematician A. Esenin-Volpin into a psychiatric clinic by force, see [4] for details related to this letter), or his ethnic origin, or unsanctioned contacts with foreigners. But even if he was not on this list, he would be obliged to pass a long and humiliating procedure with no guarantee of success. He should have got a recommendation from his local party committee and a director of the University ("kharakteristika"), then the candidature must have been approved by the special "exit commission" of the regional party committee, next, KGB permission was required, and finally he should have passed a review by the Central Party Committee of the USSR. Let me tell you a funny story about my own experience as an illustration. I was on the "black list" and had no chance to travel abroad. However, in 1976 I made an attempt. It was on the occasion of the 9th International Symposium on Mathematical

Programming in Budapest, and I was invited to present a plenary talk. Luckily, I got permission (maybe it was easier, because Hungary was in the East block). While visiting the mountains before the trip, I unfortunately broke my leg. But the desire to go was so strong that I was ready to travel in spite of the medical problems. Thus I arrived to my last appointment to get the final instructions and my passport. However, the vigilant official looked at my crutches and said: “Provide me a permission from your doctor that you can go abroad.” Well, I got it and delivered it to the bureaucrat. “No, I refuse you,” he declared. “Why?” I whispered. “What an impression will foreign scientists get when observing you? That the Soviet science stands on crutches.” And I had to stay in Moscow...

But visits abroad were not the only side of scientific life which heavily depended on the authorities’ decisions. Promotion, defense of a candidate or doctoral dissertation, elections to the Academy of Sciences, getting various awards, – everything was under the party control, and there was no chance to advance without an approval. In many cases, the lack of such an approval was explained not by individual or professional qualities, but by the data in the personal “anketa” (when hired, everybody must have filled in a detailed standard questionnaire or “anketa”). For instance, famous was “the fifth point” (item number five in the questionnaire) about your ethnic origins. If you were unlucky enough to be a Jew, your career opportunities were greatly restricted. Of course, I am unable to highlight this important theme in full detail; the interested reader can find more information in [5–7]. And many brilliant scientists, in particular in the field of mathematical programming, were forced to emigrate due to these insulting restrictions.

Another source of difficulties for researchers was the mania for secrecy. Nobody was allowed to publish any paper without special permission confirming that the publication does not contradict numerous security restrictions. All letters abroad (as well as letters from abroad) were opened and inspected. Everybody must have special permission, and a full text of the talk had to be approved if you were going to an international conference. And working in a classified institution (which was the case for many experts in mathematical programming), complicated the situation drastically.

The list of misfortunes of Soviet science looks too long... to finish, let me mention the last (but probably not the least) trouble. Most branches of science had a strongly hierarchical structure. There was “a big boss” who completely controlled all principal decisions (visits abroad, elections, awards, high level promotions, etc.). For instance, such a boss in numerical analysis was academician A.N. Tikhonov. Next, there were “local bosses” who made decisions on the local level (low level promotions, defense of candidate dissertations and so on). Luckily, there was no strong monopolization in the field of mathematical programming. Maybe the lack of a single all-powerful leader made Soviet mathematical programming more competitive and led to significant progress in this field.

Needless to say, all above-mentioned restrictions and misfortunes had their own dynamics. The situation in the 1940s to mid-1950s was the worst. Malevolent intent by the authorities could lead a researcher to the GULAG. The period 1955–1970 was the least oppressive, it was a “golden age” of Soviet mathematics [8] and it was the best time for mathematical programming as well. The years 1970 to 1985 were a period of stagnation in political, social, economic, and scientific life. All the troubles I have

mentioned above played a more and more significant role in the development of Soviet science and thus led to its degradation.

Now, after these introductory remarks oriented towards a Western or a young reader (Russian readers of my generation are too familiar with all the peculiarities of the former Soviet scientific life and could contribute a lot to my exposition based on their personal experience), we can turn to the historical perspective of the development of mathematical programming in the USSR. The roots of this development lie in the previous centuries, and we can repeat after Newton: "We stand on the shoulders of giants."

2. Prehistory

The first scientist who dealt with optimization in Russia was a true giant.

L. Euler, 1707–1783, lived in Russia in 1727–1741 and 1766–1783 (more then 30 years in total), and published about 850 papers and books. His biography and a description of his works can be found in [9]. Euler recognized the role of optimization problems; he wrote: *"Whatever happens in the world, within can be observed the meaning of maximum or minimum; hence there is no doubt that all phenomena in nature can be explained via the maximum and minimum method..."* , see [10], Appendix 1. He made important contributions to various fields of the theory and methods of optimization. First, he obtained higher order necessary and sufficient conditions of optimality for unconstrained optimization. Second, he was one of the founders of the *calculus of variations*. In this direction, he obtained a number of fundamental results such as a necessary condition for an extremum (referred to as the *Euler equation* today). Next, he applied *discrete approximations* to solving the problems of calculus of variations; this can be thought of as the first numerical method for optimization problems. Finally, Euler considered *isoperimetric problems* in the calculus of variations, treating them in a very broad sense; this can be viewed as the first studies in constrained optimization.

In the XIX-th century, a non-formal follower of Euler's research was one of the outstanding Russian mathematicians.

P.L. Chebyshev, 1821–1894. In the field of optimization, he introduced what is now known as *Chebyshev approximation*. In the simplest form, it is the problem of finding

$$\min_x \max_{t\in T} \left| a(t) - \sum_i x_i f_i(t) \right|.$$

We can say today that this is an example of a *convex nonsmooth optimization* problem, or, more specifically, *semiinfinite programming*. Chebyshev had found an analytic solution to some particular cases (for $a(t) = 1$, $f_i(t) = t^i$, $T = [0, 1]$, the solution is the *Chebyshev polynomial*) and provided a general characterization of the optimal approximation. He also studied many other optimization problems; some of them were of a practical origin (such as the least distorted geographic maps, the optimal cut out of clothes, the best choice of the parameters of mechanical devices for plotting curves). Like Euler, Chebyshev did understand the significance and diversity of extremum problems. For instance, he claimed: "*. . . the same problem is common to all practical activity*

of human beings: How to allocate our resources so as to achieve as high a profit as possible? Solution of such problems constitutes the subject of the so-called theory of maximum and minimum quantities. These problems, being of a purely practical origin, have a specific significance for the theory as well: all laws governing the movement of weighted or weightless matter are solutions of problems of this sort. We must not overlook their fruitful influence on the development of mathematical sciences."

Two great Chebyshev's followers continued his research of extremum problems.

A.A. Markov, 1856–1922, is well known for his works in number theory and probability theory (*Markov chains, Markov processes*). He also contributed to some areas of optimization. For example, he considered the so-called *problem of moments*:

$$\min \int_a^b t^n f(t)dt,$$

$$0 \le f(t) \le L, \quad \int_a^b t^i f(t)dt = c_i, \qquad i = 1, \dots, n-1.$$

This is a nonstandard constrained optimization problem with integral functionals (involving no derivatives in contrast to the calculus of variations). He also established the *Markov inequality*, which is a solution of the following problem:

$$\max_{P(x)\in\Pi} \max_{a\le x\le b} \left| P'(x) \right|$$

$$|P(x)| \le M, \quad a \le x \le b,$$

where Π is the set of all polynomials of degree at most n.

A.M. Lyapunov, 1857–1918. At first glance, his works were not related to optimization. However, this is not exactly the case. Lyapunov developed stability theory for ordinary differential equations; in its simplest form it states that a solution $x(t)$ of the equation

$$\dot{x} = f(x)$$

is stable if there exists a function $V(x)$ (the *Lyapunov function*) such that

$$\Big(\nabla V(x),\, f(x)\Big) < 0.$$

We can take a reverse point of view: The differential equation above is a continuous-time method for minimizing $V(x)$. Thus the method provides a systematic tool for validation of the convergence of numerical methods of optimization.

3. The pioneer

L.V. Kantorovich, 1912–1986. He graduated from Leningrad University at the age of 18 and became Full Professor at 22; his first paper was published when he was 16. He was granted an unusual combination of awards including the Stalin Prize (1949),

Lenin Prize (1965) and the Nobel Prize (1975). The biographical data on Kantorovich, including his autobiography *My Journey in Science*, can be found in the volume [13] published in English, see also [14–16]. Kantorovich made significant contributions to numerous areas of functional analysis; he was one of the founders of numerical analysis in our country; he was one of the first who recognized computer science as a new branch of mathematics. However, for the mathematical programming community, he is undoubtedly the father of a new science *OPTIMIZATION*, which includes the standard mathematical programming. He was responsible for the following three breakthroughs in optimization:

- Linear programming, 1939;
- General conditions of optimality, 1940;
- Functional analysis techniques, 1939–1948.

Let us consider briefly each of these.

3.1. Linear programming

In 1939, L.V. Kantorovich published a small book (just 67 pages) [17], where a new type of optimization problem was addressed. It was not in the standard form of linear programming, but used a number of variations. Much later, in [20], Kantorovich stated that the model considered in the Western literature was a particular case of his model; the equivalence of all these formulations of a linear programming problem was not clearly understood at that time. The appearance of this book was stimulated by the questions of engineers and economists and was not based on any previous mathematical studies of the author or other mathematicians. Many practical applications were indicated in the book; also, ideas of numerical methods based on dual variables (called "resolving factors") were given.

However, this revolutionary book had little response from economists or mathematicians! There were several reasons for such disregard. First, there was no true demand for mathematical methods in the totalitarian system of the Soviet Union. Contrary to the first phrase in [20] *"Planning of national economy or its individual branches on the scale of the state is only possible under the condition that the private (capitalist) ownership of the means of production is replaced with the peoples' (socialist) ownership,"* the socialist system was based on voluntary, administrative methods. The ideas of "higher rationality" (i.e., what was favorable to the leaders of the country at the moment) prevailed over common sense and rational reasonings (see Sect. 5.1.7 below). Second, the book was not written as a mathematical text, that is why mathematicians paid no attention to it. It was just a collection of examples, word explanations and simple calculations. The first rigorous mathematical paper on the general linear programming problem was published by Kantorovich only in 1957. His second book on linear programming [20] was published in 1961. It contained two mathematical appendices (written in collaboration with G. Rubinstein) where a mathematical formulation of linear programming and numerical methods for its solution were given; however, the main effort was directed to the provision of compatibility with the Marxist dogma. For instance, dual variables were referred to as *"objective-conditioned estimates"* ("o.o.o." in Russian) rather than "prices."

3.2. General conditions of optimality

In [18], Kantorovich considered a very general constrained optimization problem in a topological space:

$$\min_{x \in Q} f(x).$$

Here, $f(x)$ is a differentiable functional in a topological space X, $Q \subset X$ is a set which yields a convex cone approximation K at a point $x^* \in Q$. He obtained the following necessary condition for extremum.

Theorem. *If x^* is the minimum point, then $f'(x^*) \in K^*$.*

In the theorem, $f'(x)$ denotes the gradient of $f(x)$, while K^* is the cone conjugate to K. Surprisingly, this brilliant mathematical paper also caused no response! Probably because it was, in fact, ahead of its time.

3.3. Functional analysis techniques

The first paper on the convergence of an iterative method for minimization of a quadratic functional was published by Kantorovich in 1939. Later (1944, 1945, 1947), he continued this research. The results were summarized in the large paper [19], which was published in 1948 in the leading Soviet mathematical journal *Russian Mathematical Surveys*. Among many other results related to numerical analysis, it contained the proofs of convergence (and estimates of the rate of convergence) of the steepest descent method for a quadratic functional and Newton's method for functional equations. Besides these practical results, the paper demonstrated successful applications of functional analysis techniques to optimization problems. Concluding, this work established a high level of standards for the validation of numerical methods in optimization.

4. Fifties: emerging a new science

By the 1950s, a number of research directions in the field of extremum problems had appeared.

Several works dealt with *linear inequalities, Chebyshev approximations to inconsistent linear equations* and related topics. S.N. Chernikov studied the theory of *linear inequalities* ("the principle of bounding solutions" which describes the vertices of the solution set). Two scientists from the Ukraine, E.Ya. Remez and S.I. Zukhovitskii proposed *finite methods for finding the best Chebyshev approximation*; actually, these methods were simplex-like algorithms for solving the corresponding linear programming problems (primal or dual). This research is summarized in the monographs [21–23] published much later than the original works. For instance, the first book (in Ukrainian) on numerical methods for Chebyshev approximations was written in 1935.

At the same time, significant progress was achieved in optimization problems related to *Markov's problem of moments*, due to the scientific schools of N.I. Akhiezer and M.G. Krein [24, 25].

Simultaneously intensive research was initiated in the field of *linear programming*. G.Sh. Rubinstein, a former student of L. Kantorovich published the first rigorous mathematical formulation of linear programming problems and their analysis in Russian in 1955 [26]. The results of Western scientists (G. Dantzig, H. Kuhn, A. Tucker, D. Gale, and others) on linear programming had became known in the Soviet Union. An important role in this process was played by translations of the original works into Russian; the first one was [27]. Russian textbooks on linear programming and related topics appeared. The first publications on matrix games were by N.N. Vorobjev and E.S. Ventzel [29], while the first Russian textbook on linear programming was written by D.B. Yudin and E.G. Golstein [28].

Important events happened in the field of automatic control. In 1956, L.S. Pontryagin and his coauthors provided a new formulation of the *optimal control* problem and obtained a new necessary condition for optimality, the so-called *maximum principle*. It was a far-reaching extension of results in the calculus of variations [30].

As the reader can conclude, it was a time of intensive activity. However, the results obtained were isolated and were not understood as being parts of a unified scientific discipline. Are there any links between linear programming and optimal control, between the problem of moments and Chebyshev approximations,and between numerical methods for unconstrained optimization and those for solving linear inequalities? In the 1950s, the typical answer was negative.

5. The sixties: golden age

The situation had changed dramatically by the early 1960s. The time was ripe; all parts of the puzzle were ready to deliver a beautiful overall picture, resulting in an immediate kaleidoscope of new contributions.

5.1. Main research directions

5.1.1. The general theory of extremum problems. The understanding of the common nature of various optimization problems was the first breakthrough of the period. In the paper [18] by Kantorovich, the first general framework for the analysis of extremum problems was formulated; however it provided no technique for the specification of general conditions of extremum for particular problems. The *Dubovitskii-Miljutin formalism* [31] succeeded in this direction: a theorem on the conjugate cone for an intersection of several cones was an efficient tool for formulating necessary conditions of the extremum in a unified manner. This approach was well suited for a broad class of optimization problems mentioned above, such as linear programming, optimal control, approximation theory and many others. Moreover, it provided new optimality conditions for some hard problems (for instance, optimal control subject to phase constraints). The technique became very popular among Soviet researchers; the book [33] by I. Girsanov played a significant role in this popularization. Later, V. Boltyanskii [32] extended the approach and called it *the method of tents*.

Another approach, developed by B. Pshenichnyj in the mid-sixties (and summarized in his books [34, 35]), was based on the methods of *convex analysis*. It also exploited

the tools of functional analysis for obtaining necessary and sufficient conditions of optimality. The fundamental monograph by A. Ioffe and V. Tikhomirov continued this line of research [37]. Similar studies in the West were carried out by R.T. Rockafellar, L. Neustadt, H. Halkin, L. Berkovitz, J. Varga and others.

General *duality theory* for convex optimization problems was developed by E. Golstein [36].

5.1.2. Numerical methods for general extremum problems. In parallel with the development of the general theory of extremum problems it was realized that the numerical methods for solving them can also be treated in a unified framework. The publications [38,39] introduced numerous gradient-like methods for unconstrained optimization and their extensions to constrained optimization such as *gradient projection, conditional gradient, constrained Newton, cutting plane, penalty function methods* and some others. The general theorems on convergence and the rate of convergence of such methods in finite and infinite dimensional problems were proved; numerous applications (to standard mathematical programming problems, optimal control, semiinfinite programming, etc.) were considered. This line of research remained very active for some time, e.g., see the monographs [40–42].

5.1.3. Nonsmooth optimization. Traditionally, optimization methods dealt with differentiable functions and were based on gradient approximation. N. Shor was the first who extended the approach to nonsmooth convex optimization. In his Ph.D. thesis (1964), he proposed a subgradient method for nondifferentiable functions and applied this approach to numerical solution of a program dual to a transportation-like problem of linear programming. Later, this approach was extended and validated by N. Shor, Yu. Ermoliev, and B. Polyak [44–46]. The monograph [47] summarizes these studies; see also, [43].

A different method, the so-called *center of gravity method* was proposed by A. Levin [48] (and independently by Newman in the USA). Later it was proved to be optimal, with respect to the order, amongst all the methods of nonsmooth convex optimization which exploit only the values of the subgradient; however, it included the operation of finding the center of gravity of a polytope, which is hard to perform.

Numerous methods have been developed for particular classes of nonsmooth optimization, primarily for *minimax problems*, by V. Demyanov [49,50].

5.1.4. Stochastic optimization. In many cases, the values of functions and their gradients are corrupted by random noise. Effective minimization methods should be modified to retain convergence in such situations. This can be done via averaging or by step regulation in iterative methods similar to the approach used by *stochastic approximation* methods in statistics. Such a revision of minimization methods under random noise was undertaken by Yu. Ermoliev [51]. Numerous applications of iterative stochastic algorithms to the problems of identification, estimation, pattern recognition, etc., can be found in the book by Ya. Tsypkin [52].

Sometimes randomness is artificially incorporated in the minimization process, as for example in *random search* methods. An active proponent of such methods was L. Rastrigin [53].

There was not that much activity in the area of *stochastic programming* (in the sense this term is understood in the Western literature); however a textbook was written by D. Yudin [54].

5.1.5. Linear programming and related topics. After a long pause, investigations in this area resumed in the USSR. They were focused on the development of *software* for implementations of the *simplex method* and its variations (I. Romanovskii, U. Malkov, K. Kim, A. Cherkasski, V. Skokov, A. Stanevichus and others), numerical methods for special classes of linear programming problems (*transportation problems, decomposition methods, Chebyshev approximations*) [55], and *iterative methods* for linear programming [56–58]. The fate of the paper [56] was especially interesting. I. Dikin was a student of Kantorovich, and in his Ph.D. thesis he formulated rigorously some heuristic rules which his supervisor suggested for the numerical solution of linear programming problems. Dikin proved the convergence of the iterative algorithm, but failed to estimate its rate of convergence. The paper (similarly to the earlier works of Kantorovich, see above) attracted no attention and had been forgotten till the late 1980s, when it was recognized that the implemented version of the famous Karmarkar algorithm [65] was very close to Dikin's original method.

5.1.6. Discrete optimization. In Russian, the first monograph on discrete programming was published in 1969 [59]. The efforts of researchers in this field (Yu. Finkelstein, A. Korbut, A. Fridman, E. Levner, I. Sigal, I. Sergienko, V. Emelichev, A. Karzanov, E. Dinits, S. Lebedev and others) were mostly directed towards the solution of a special classes of combinatorial problems.

5.1.7. Applications. In the early 1960s there were very optimistic beliefs about the practical applications of the ideas of optimal planning to the socialist economy and society. Many mathematicians were convinced that optimal plans and prices can be calculated using linear programming models. However, real-life attempts to implement this approach had failed. Many stories on the subject can be found in [16]. Here is just one of them. Exploiting Kantorovich's work on the optimal cutting of a rectangular sheet into pieces of a given shape, engineers and economists of a factory that manufactured steel items were able to increase considerably the output production. However, they faced quite unexpected obstacles. First, as an outcome, the plan for the next year had to be increased correspondingly (for the socialist planning system, it was customary to require some extension of production automatically each year), but now the factory had no reserves to fulfill the expanded future plan. Second, each enterprise had a plan for collecting scrap. Obviously, due to the optimal cutting strategy, the amount of unused steel decreased, and thus this plan was violated. As a result, top managers of the factory were very upset (they were reprimanded by the Communist party officials) and consequently they refused to collaborate with mathematicians in future.

Much later, Kantorovich organized a special commission under his guidance to investigate the situation with respect to applications of optimization models to the Soviet economy. The main goal was to find examples of successful applications and to extend the experience to other branches of production and service. After considerable investigation, the commission recognized the complete lack of such positive experiences.

5.2. Main scientific centers

In contrast to the vast collection of research communities in the USA, where it is spread over many universities and some government and industrial research institutes, the mathematical programming studies in the USSR were concentrated in a few scientific centers:

Moscow: Moscow State University (I. Girsanov, V. Tikhomirov, B. Polyak, F. Vasiliev), Yudin's Laboratory in a classified research institute (D. Yudin, E. Golstein, later A. Ioffe, A. Nemirovskii), Central Economical-Mathematical Institute (E. Golstein, later V. Skokov, N. Tretyakov, Yu. Nesterov), Computer Center of the Academy of Sciences (N. Moiseev, Yu. Evtushenko, later L. Khachiyan, A. Antipin), other institutes (A. Dubovitskii, A. Miljutin);

Kiev: Glushkov Institute of Cybernetics (Yu. Ermoliev, B. Pshenichnyj, N. Shor, V. Mikhalevich, E. Nurminskii), other institutes (S. Zukhovitskii, R. Polyak, M. Primak);

Novosibirsk: Institute of Mathematics and Novosibirsk State University (L. Kantorovich, G. Rubinstein, I. Dikin, V. Bulavskii, A. Rubinov, A. Kaplan);

Leningrad: Leningrad State University (V. Demyanov, I. Romanovskii, A. Vershik).

Besides these four centers, research activity took place in some provincial towns such as Kharkov (Yu. Ljubich, G. Majstrovskii), Sverdlovsk (I. Eremin), Irkutsk (V. Bulatov), Minsk (R. Gabasov, F. Kirillova, later B. Mordukhovich), Voronezh (M. Krasnoselskii, A. Levin), etc.

5.3. Main journals

Publications on mathematical programming appeared in various journals. The main journals with their English translations and places of publication are listed below.

- Zhurn. Vychisl. Matem. i Matem. Fiz. = USSR Journ. Comp. Math. and Math Phys., Moscow; a journal on numerical analysis.
- Kibernetika = Cybernetics, Kiev; a journal on computer science, system theory and optimization.
- Ekonomika i Matematicheskie Metody = Mathecon, Moscow; a journal on mathematical economics.
- Avtomatika i Telemekhanika = Automation and Remote Control, Moscow; a journal on control.
- Doklady AN SSSR = Soviet Math. Doklady, Moscow; all branches of science.

However, we have never had a special journal on mathematical programming!

5.4. Main events

The life of the optimization community in the 1960s was full of professional events, to name just a few of the most important ones:

- *Seminars and lecture courses on optimization in Moscow State University*, started 1961, I. Girsanov, V. Tikhomirov, B. Polyak;
- *Seminars and lectures in Kiev*, started 1960, S. Zukhovitskii, B. Pshenychnyj, N. Shor;
- *International Congress of Mathematicians*, Moscow, 1966. At the time this was the first contact with Western experts, and so played a significant role in our involvement in the international community;
- *Drogobych Winter Schools on Mathematical Programming and Related Topics.* Chaired by S. Zukhovitskii and held annually since 1968. Some years, there were as many as 500(!) participants;
- *Summer Schools on Optimization*, chaired by N. Moiseeev, started 1966;
- *All-Union Symposia on Optimal Programming Software*, chaired by E. Golstein, started 1970;
- *All-Union Conferences on Mathematical Programming*;
- *All-Union Symposia on Extremum Problems*, started 1963;
- *All-Union Conference on Numerical Analysis*, 1965, Moscow, had many sessions devoted to optimization problems.

6. The 1970–80s: new directions

As I have mentioned above, the period after the 1970s was characterized by a progressive decline of many research areas in the USSR; this had an effect on mathematical programming as well. There were many internal reasons for this (specific to the development of any science) as well as the general atmosphere of stagnation in Soviet society, which played its unavoidable role. In any case, activity of the previous community life decreased (fewer meetings on optimization with fewer attendees, fewer new ideas etc.) Nevertheless, there were several breakthroughs during this period, and I will mention some of them below.

6.1. Complexity of optimization problems and efficient methods of optimization

In the series of publications of 1976–1979 (summarized in the monograph [60]), A. Nemirovskii and D. Yudin introduced a new notion of *complexity of optimization problems*. Their approach was as follows. Let us consider a family of optimization problems equipped with an *oracle*, i.e., a certain source of information about one or another particular element in the family we deal with. For instance, suppose we consider a class of smooth, strongly convex unconstrained minimization problems, then the oracle provides a value of a function to be minimized and its gradient at any point. What are the potential abilities of arbitrary methods which exploit this information? Nemirovskii and Yudin established lower bounds for various classes of optimization problems. For instance, they have found that for the above-mentioned class of strongly convex smooth optimization problems, there exists no method which provides a solution with a relative error ν for less than $O(\sqrt{Q}\ln 1/\nu)$ calculations of the gradient, where Q is the ratio of the Lipshitz constant for the gradient and the constant of strong convexity. Similar

results were obtained for the classes of Lipshitz continuous (multiextremum) problems, general convex problems, stochastic optimization problems (where the oracle provides a gradient value corrupted by random noise), etc.

On this way, the authors were able to find *effective methods* for minimization. Indeed, if a method has complexity which coincides by the order with the lower bound, then it is the optimal one, i.e., there exists no other method that solves all problems of the class in a faster (by the order) way. Note that the notion of complexity in [60] differs from the standard asymptotic rate of convergence – it is not an asymptotic property.

A serious contribution to the theory of efficient methods was made by Yu. Nesterov [61]. For instance, he has found the optimal method for minimization of smooth convex (but not strongly convex) functions.

6.2. The method of ellipsoids and polynomial complexity of linear programming

Above, I have mentioned the "center of gravity" method by A. Levin [48] for the minimization of nondifferentiable convex functions. An implementable version of this method (where an auxiliary problem of finding the center of gravity of a polytope was replaced with the trivial problem of finding the center of an outer ellipsoid) was proposed by Yudin and Nemirovskii in 1976 [62] and independently, a year later, by N. Shor [63]. Based on this method of ellipsoids, L. Khachyan constructed an iterative method for solving linear programming problems which was proven to have *polynomial complexity* [64]. Hence, the question: Whether linear programming problems are NP-hard or not, which remained open for a long time, has been resolved.

6.3. Interior-point polynomial methods and semidefinite programming

In spite of being solid theoretically from the complexity point of view, the method of ellipsoids for linear programming had not become a competitor to the simplex method. The true numerical progress is connected with Karmarkar's method [65]. Following this line of research, Yu. Nesterov and A. Nemirovskii considerably extended the approach in the series of papers published in 1987–1989 and summarized later in their monograph [66]. They constructed *polynomial interior-point methods* for various classes of convex programming problems. Moreover, they developed general notions and techniques (*self-concordant functions, barriers*) for such an extension. One of their significant contributions was the development of *semidefinite programming*, i.e. optimization problems subjected to matrix semidefiniteness constraints. They obtained polynomial algorithms for solving such problems based on interior-point methods with self-concordant barriers. These pioneering works generated a new area of mathematical programming [67]. Moreover, a particular class of semidefinite programming, so-called *linear matrix inequalities*, has found numerous applications in control [68].

6.4. Progress in nonconvex and nonsmooth analysis

The general analysis techniques developed in the 1960s for optimization problems were based either on the classical calculus (for smooth problems) or on convex analysis

(for nonsmooth convex problems). New types of application required the development of more sophisticated tools which have got the name *"nonconvex analysis."* One of the most successful approaches to nonconvex analysis was proposed by B. Mordukhovich [69]; it is based on the *method of metric approximations*. Another way was chosen by A. Ioffe (e.g., see [70]); it exploits the extension of the notion of *subdifferential* to nondifferentiable mappings. Yu. Nesterov [61] elaborated the theory of *lexicographic differentiation* which allowed to obtain calculus of nonsmooth functions. E. Levitin systematically studied the *theory of perturbations* for smooth and nonsmooth optimization problems [71].

7. Conclusion

I am writing the paper after many dramatic events in our country, such as the collapse of the USSR and the end of the communist system. Of course, political and economical changes strongly affected the situation with research in Russia. The organization of basic research was not modified and looks outdated. The shortage of state funding and the lack of support from industry and business unavoidably leads to slow but permanent decline of research institutions. Since there are no former obstacles for going abroad, many experts choose emigration as a way to resolve their personal problems. Today, a minor part of the mathematical programming community remains in the country. Very few seminars, conferences, workshops are held in Russia, and researchers face strong difficulties in visiting such meetings abroad due to the lack of financing. Last but not least, – research does not attract young people any more, and the community is getting older.

Nevertheless, I do not want to finish on such a pessimistic note. I do believe that the great traditions of mathematical research in Russia will be able to overcome the difficulties, and science will survive...

Acknowledgements. The author thanks his numerous colleagues and friends for helpful discussions on the subject of the paper. The assistance of P. Shcherbakov and A. Conn in preparation of the text is gratefully acknowledged.

References

1. Karmanov, V.G. (1975): Mathematical Programming. Moscow, Nauka, 272 pp.
2. Tikhomirov, V.M. (1986): Stories on Maxima and Minima. Moscow, Nauka, 190 pp.
3. Demidov, S.S., Levshin, B.V., eds. (1999): The Case of Academician N.N. Luzin. St. Petersburg, RHGI, 310 pp.
4. Esenin-Volpin, A.S. (1999): Philosophy, Logic, Poetry, Protection of Human Rights. Moscow, RGGU, 450 pp.
5. Freiman, G. (1980): It Seems, I am a Jew. South. Illinois Univ. Press, Feffer & Simons, Inc., 97 pp.
6. Vershik, A. (1994): Admission to the mathematical faculty in Russia in the 1970s and 1980s. Math. Intelligencer **16**(4), 4–5
7. Shen, A. (1994): Entrance examinations to the Mekh-mat. Math. Intelligencer **16**(4), 6–10
8. Zdravkovska, S., Duren, P.L., eds. (1993): Golden years of Moscow Mathematics. Amer. Math. Soc. and London Math. Soc.
9. Yushkevich, A.P. (1968): History of Mathematics in Russia before 1917. Moscow, Nauka, 592 pp.

10. Euler, L. (1744): Methodus inveniendi lineas curvas maximi minimive propietate gaudentes, sive solutio problematis isoperimetrici latissimo sensu accepti. Lausannae et Genevae (in Latin). Russian translation: Euler, L. (1934): A Method for Finding Curves Having the Property of Maximum or Minimum, or a Solution of the Isoperimetric Problem Considered in the Very Broadest Sense. Moscow-Leningrad, Academy of Sciences
11. Mathematics in the USSR for Forty Years, 1917–1957. Moscow, GIFML, 1959, Vol. 1, 1002 pp., Vol. 2, 819 pp.
12. Mathematics in the USSR, 1958-1967. Moscow, GIFML, 1969, Vols. 1–2, 1579 pp.
13. Leifman, L.J., ed. (1990): Functional Analysis, Optimization, and Mathematical Economics: A Collection of Papers Dedicated to the Memory of Leonid Vital'evich Kantorovich. Oxford University Press
14. Romanovskii, I.V. (1989): L.V. Kantorovich's works in mathematical programming. In: Iri, M., Tanabe, K., eds., Mathematical Programming: Recent Developments and Applications. Tokyo, Kluwer, pp. 365–382
15. Kantorovich, L.V. (1987): My way in science. Uspekhi Matem. Nauk **42**(2), 183–213
16. Katzenelinboigen, A., Kantorovich, L.V. (1997): A political dilemma in scientific activity. Ekonom. i Matem. Metody **33**(3), 30–42
17. Kantorovich, L.V. (1939): Mathematical Methods for Production Organization and Planning. Leningrad, 64 pp.
18. Kantorovich, L.V. (1940): On an efficient method for solving some classes of extremum problems. Doklady AN SSSR **28**, 212–215
19. Kantorovich, L.V. (1948): Functional analysis and applied mathematics. Uspekhi Matem. Nauk **6**(3), 89–185
20. Kantorovich, L.V. (1960): Economical Calculation of the Best Utilization of Resources. Moscow, Academy of Sciences, 347 pp.
21. Chernikov, S.N. (1968): Linear Inequalities. Moscow, Nauka, 488 pp.
22. Remez, E.Ya. (1969): Foundations of Numerical Methods for Chebyshev Approximation. Kiev, Naukova Dumka, 623 pp.
23. Zukhovitskii, S.I. (1956): On the approximation of real functions in the sense of P.L. Chebyshev. Uspekhi Matem. Nauk **11**(2), 125–159
24. Akhiezer, N.I. (1961): The Classical Problem of Moments. Moscow, Fizmatgiz, 310 pp.
25. Krein, M.G., Nudelman, A.A. (1973): Markov's Problem of Moments and Extremum Problems. Moscow, Nauka, 551 pp.
26. Rubinstein, G.Sh. (1955): Problem on the extreme point of intersection of an axis with a polytope. Dokl. AN SSSR **115**(3), 627–630
27. Linejnye neravenstva i smejnye voprosy. (pod red. H. Kuhn i A. Tucker) Moscow, IL, 1959; Russian translation of: Linear Inequalities and Related Systems. Kuhn, H., Tucker, A., eds., Ann. Math. Studies, No. 38, Princeton, Princeton Univ. Press, 1956
28. Yudin, D.B., Golstein, E.G. (1961): Problems and Methods of Linear Programming. Moscow, Sov. Radio, 491 pp.
29. Ventzel, E.S. (1959): Elements of Game Theory. Moscow, Fizmatgiz
30. Pontryagin, L.S., Boltyanskii, V.G., Gamkrelidze, R.V., Mischenko, E.F. (1961): Mathematical Theory of Optimal Processes. Moscow, Nauka
31. Dubovitskii, A.Ya., Miljutin, A.A. (1963): Extremum problems under constraints. Dokl. AN SSSR **149**(4), 759–762
32. Boltyanskii, V.G. (1975): Method of tents in the theory of extremum problems. Uspekhi Matem. Nauk **30**(3), 3–55
33. Girsanov, I.V. (1970): Lectures on the Mathematical Theory of Extremum Problems. Moscow, Moscow State Univ., 118 pp.
34. Pshenichnyj, B.N. (1969): Necessary Conditions for Extrema. Moscow, Nauka, 152 pp.
35. Pshenichnyj, B.N. (1980): Convex Analysis and Extremum Problems. Moscow, Nauka, 320 pp.
36. Golstein, E.G. (1971): Duality Theory in Mathematical Programming and Its Applications. Moscow, Nauka, 352 pp.
37. Ioffe, A.D., Tikhomirov, V.M. (1974): Theory of Extremum Problems. Moscow, Nauka, 479 pp.
38. Levitin, E.S., Polyak, B.T. (1966): Minimization methods under constraints. Zhurn. Vychisl. Matem. i Matem. Fiz. **6**(5), 787–823
39. Demyanov, V.F., Rubinov, A.M. (1968): Approximate Methods for Solving Extremum Problems. Leningrad, Leningr. State Univ., 180 pp.
40. Pshenichnyj, B.N., Danilin, Yu.M. (1957): Numerical Methods in Extremum Problems. Moscow, Nauka, 320 pp.
41. Vasiliev, F.P. (1974): Lectures on the Methods for Solving Extremum Problems. Moscow, Moscow State Univ., 374 pp.

42. Polyak, B.T. (1983): Introduction to Optimization. Moscow, Nauka, 384 pp.
43. Shor, N.Z. (1991): The development of numerical methods for nonsmooth optimization in the USSR. In: Lenstra, J.K., Rinnoy Kan, A.H.G., Shrijver, A., eds., History of Mathematical Programming. Amsterdam, CWI, North-Holland, pp. 135–139
44. Ermoliev, Yu. (1966): Methods for solving nonlinear extremum problems. Kibernetika **4**, 1–17
45. Polyak, B.T. (1967): A general method for solving extremum problems. Dokl. AN SSSR **174**(1), 33–36
46. Polyak, B.T. (1969): Minimization of nonsmooth functionals. Zhurn. Vychisl. Matem. i Matem. Fiz. **9**(3), 509–521
47. Shor, N.Z. (1979): Minimization Methods for Nondifferentiable Functions and Their Applications. Kiev, Naukova Dumka, 200 pp.
48. Levin, A.Yu. (1965): On an algorithm for minimization of convex functions. Dokl. AN SSSR **160**(6), 1244–1247
49. Demyanov, V.F., Malozemov, V.N. (1972): Introduction to Minimax. Moscow, Nauka, 368 pp.
50. Demyanov, V.F., Vasiliev, P.V. (1981): Nondifferentiable Optimization. Moscow, Nauka, 384 pp.
51. Ermoliev, Yu.M. (1976): Stochastic Programming Methods. Moscow, Nauka, 240 pp.
52. Tsypkin, Ya.Z. (1968): Adaptation and Learning in Automatic Systems. Moscow, Nauka, 400 pp.
53. Rastrigin, L.A. (1968): Statistical Methods for Finding Extrema. Moscow, Nauka, 376 pp.
54. Yudin, D.B. (1979): Problems and Methods of Stochastic Programming. Moscow, Sov. Radio
55. Golstein, E.G., Yudin, D.B. (1966): New Directions in Linear Programming. Moscow, Sov. Radio, 524 pp.
56. Dikin, I. (1967): An iterative solution of problems of linear and quadratic programming. Dokl. AN SSSR **174**(4), 747–748
57. Polyak, B.T., Tretyakov, N.V. (1972): On an iterative method of linear programming and its economic interpretation. Econom. i Matem. Metody **8**(5), 740–751
58. Belenkii, V.Z., Volkonskii, V.A., a.o. (1974): Iterative Methods in Games Theory and Programming. Moscow, Nauka
59. Korbut, A.A., Finkelstein, Yu.Yu. (1969): Discrete Programming. Moscow, Nauka
60. Nemirovskii, A.S., Yudin, D.B. (1979). Complexity of Problems and Efficiency of Optimization Methods. Moscow, Nauka, 383 pp.
61. Nesterov, Yu.E. (1989): Effective Methods in Nonlinear Programming. Moscow, Radio i Svyaz, 301 pp.
62. Yudin, D.B., Nemirovskii, A.S. (1976): Information complexity and effective methods for solving convex extremum problems. Ekonom. i Matem. Metody **12**(2), 357–369
63. Shor, N.Z. (1977): A cutting method with space dilation for solving convex programming problems. Kibernetika **1**, 42–50
64. Khachyan, L.G. (1979): Polynomial methods in linear programming. Dokl. AN SSSR **244**(5), 1093–1096
65. Karmarkar, N. (1984): A new polynomial-time algorithm for linear programming. Combinatorica **4**, 373–395
66. Nesterov, Yu.E., Nemirovskii, A.S. (1994): Interior-Point Polynomial Algorithms in Convex Programming. Philadelphia, SIAM, 405 pp.
67. Saigal, R., Vanderberghe, L., Wolkowitz, H., eds. (2000): Handbook of Semidefinite Programming. Waterloo, Kluwer
68. Boyd, S., El Ghaoui, L., Feron, E., Balakrishnan, V. (1994): Linear Matrix Inequalities in Systems and Control Theory. Philadelphia, SIAM, 193 pp.
69. Mordukhovich, B.Sh. (1988): Approximation Methods in Optimization and Control Problems. Moscow, Nauka, 359 pp.
70. Ioffe, A.D. (1981): Nonsmooth analysis: differential calculus of nondifferentiable mapping. Trans. Amer. Math. Soc. **266**(1), 1–56
71. Levitin, E.S. (1992): Perturbation Theory in Mathematical Programming and Its Applications. Moscow, Nauka, 360 pp.